Probability and Random Variables

Saravanan Vijayakumaran
sarva@ee.iitb.ac.in

Department of Electrical Engineering
Indian Institute of Technology Bombay

July 27, 2012

Basic Probability

Sample Space

Definition (Sample Space)

The set of all possible outcomes of an experiment.

Example (Coin Toss)

$S = \{\text{Heads}, \text{Tails}\}$

Example (Die Roll)

$S = \{1, 2, 3, 4, 5, 6\}$

Example (Life Expectancy)

$S = [0, 120]$ years

Event

Definition (Event)

A subset of a sample space.

Example (Coin Toss)

$E = \{\text{Heads}\}$

Example (Die Roll)

$E = \{2, 4, 6\}$

Example (Life Expectancy)

$E = [0, 50]$ years

Definition (Mutually Exclusive Events)

Events E and F are said to be mutually exclusive if $E \cap F = \phi$.

Probability Measure

Definition

A mapping P on the event space which satisfies

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. For any sequence of events $E_1, E_2, \ldots$ that are pairwise mutually exclusive, i.e. $E_n \cap E_m = \phi$ for $n \neq m$,

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n)$$

Example (Coin Toss)

$S = \{\text{Heads}, \text{Tails}\}$, $P(\{\text{Heads}\}) = P(\{\text{Tails}\}) = \frac{1}{2}$

More Definitions

Definition (Independent Events)

Events E and F are independent if $P(E \cap F) = P(E)P(F)$

Definition (Conditional Probability)

The conditional probability of E given F is defined as

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

assuming $P(F) > 0$.

Theorem (Law of Total Probability)

For events E and F,

$$P(E) = P(E \cap F) + P(E \cap F^c) = P(E|F)P(F) + P(E|F^c)P(F^c).$$

Bayes' Theorem

Theorem

For events E and F,

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

Remarks

- Useful when $P(E|F)$ is easier to calculate than $P(F|E)$.
- Denominator is typically expanded using the law of total probability

$$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$$

Random Variables

Random Variable

Definition

A real-valued function defined on a sample space.

Example (Coin Toss)

$X = 1$ if outcome is Heads and $X = 0$ if outcome is Tails.

Example (Rolling Two Dice)

$S = \{(i,j) : 1 \leq i,j \leq 6\}$, $X = i + j$.

Cumulative distribution function

Definition

The cdf F of a random variable X is defined for any real number a by

$$F(a) = P(X \leq a).$$

Properties

- $F(a)$ is a nondecreasing function of a
- $F(\infty) = 1$
- $F(-\infty) = 0$

Discrete Random Variable

Definition (Discrete Random Variable)

A random variable whose range is finite or countable.

Definition (Probability Mass Function)

For a discrete RV, we define the probability mass function $p(a)$ as

$$p(a) = P[X = a]$$

Properties

- If X takes values $x_1, x_2, \ldots$, then $\sum_{i=1}^{\infty} p(x_i) = 1$
- $F(a) = \sum_{x_i \leq a} p(x_i)$

The Bernoulli Random Variable

Definition

A discrete random variable X whose probability mass function is given by

$$\begin{aligned} P(X = 0) &= 1 - q \\ P(X = 1) &= q \end{aligned}$$

where $0 \leq q \leq 1$.

Used to model experiments whose outcomes are either a success or a failure

The Binomial Random Variable

Definition

A discrete random variable X whose probability mass function is given by

$$P(X = i) = \binom{n}{i} q^i (1 - q)^{n-i}, \quad i = 0, 1, 2, \ldots n.$$

where $0 \leq q \leq 1$.

Used to model n independent Bernoulli trials

Continuous Random Variable

Definition (Continuous Random Variable)

A random variable whose cdf is differentiable.

Example (Uniform Random Variable)

A continuous random variable X on the interval $[a, b]$ with pdf

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}$$

Probability Density Function

Definition (Probability Density Function)

For a continuous RV, we define the probability density function to be

$$f(x) = \frac{dF(x)}{dx}$$

Properties

- $F(a) = \int_{-\infty}^{a} f(x)\, dx$
- $P(a \leq X \leq b) = \int_{a}^{b} f(x)\, dx$
- $\int_{-\infty}^{\infty} f(x)\, dx = 1$
- $P\left(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}\right) = \int_{a-\frac{\epsilon}{2}}^{a+\frac{\epsilon}{2}} f(x)\, dx \approx \epsilon f(a)$

Mean and Variance

- The expectation of a function g of a random variable X is given by

$$E[g(X)] = \sum_{x:p(x)>0} g(x)p(x) \quad \text{(Discrete case)}$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx \quad \text{(Continuous case)}$$

- Mean = $E[X]$
- Variance = $E\left[(X - E[X])^2\right]$

Random Vectors

Random Vectors

Definition (Random Vector)

A vector of random variables

Definition (Joint Distribution)

For a random vector $\mathbf{X} = (X_1, \ldots, X_n)^T$, the joint cdf is defined as

$$F(\mathbf{x}) = F(x_1, \ldots, x_n) = P\left[X_1 \leq x_1, \ldots, X_n \leq x_n\right].$$

Remarks

- For continuous random vectors, the joint pdf is obtained by taking partial derivatives
- For discrete random vectors, the joint pmf is given by

$$p(\mathbf{x}) = p(x_1, \ldots, x_n) = P\left[X_1 = x_1, \ldots, X_n = x_n\right]$$

Mean Vector and Covariance Matrix

For a $n \times 1$ random vector $\mathbf{X} = (X_1, \ldots, X_n)^T$

- Mean is

$$\mathbf{m}_X = E[\mathbf{X}] = \begin{pmatrix} E[X_1] \\ \vdots \\ E[X_n] \end{pmatrix}$$

- Covariance is

$$\begin{aligned} \mathbf{C}_X &= E\left[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^T\right] \\ &= E\left[\mathbf{X}\mathbf{X}^T\right] - E[\mathbf{X}](E[\mathbf{X}])^T \end{aligned}$$

Marginal Densities from Joint Densities

- Continuous case

$$f(x_1) = \int \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_2 \ldots dx_n$$

- Discrete case

$$p(x_1) = \sum_{x_2} \cdots \sum_{x_n} p(x_1, x_2, \ldots, x_n)$$

Bayes' Theorem for Conditional Densities

Definition (Conditional Density)

The conditional density of Y given X is defined as

$$f(y|x) = \frac{f(x,y)}{f(x)}$$

for x such that $f(x) > 0$.

Theorem (Bayes' Theorem)

$$f(x|y) = \frac{f(y|x)f(x)}{f(y)} = \frac{f(y|x)f(x)}{\int f(y|x)f(x)\,dx} \quad \textit{(Continuous)}$$

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\sum_x p(y|x)p(x)} \quad \textit{(Discrete)}$$

Thanks for your attention